TECHNISCHE UNIVERSITÄT DRESDEN

Dresden Database Research Group

Prolog

Scalable Data Engineering

Version: Monday, October 10, 2022

Who for Whom???

Dresden Database Research Group

Lecturers

- Dr.-Ing. Claudio Hartmann claudio.hartmann@tu-dresden.de
- Dipl.-Inf. Lucas Woltmann lucas.woltmann@tu-dresden.de

Lecture

- Scalable Data Engineering

CMS-CE-EL1 , CMS-CE-EL2 , CMS-CLS-ELG , CMS-CMA-ELG , CMS-CMA-ELV1 , CMS-COR-DBM , CMS-EE-EL1 , CMS-EE-EL2 , CMS-VC-ELV1 , CMS-VC-ELV2 , D-WW-INF-3421 , D-WW-INF-3422 , D-WW-INF-3423 , INF-BAS2 , INF-BAS4 , INF-BI-2 , INF-LE-WW , INF-PM-ANW , INF-VERT2 , INF-VERT4 , MINF-04-FG-IAS , WI-MA-08-01 , WI-MA-09-01

- See Lecture Catalog: https://wwwdek.inf.tu-dresden.de/mole-web/catalogs/wise2223/institute/sya/courses/de#db

Extent

- 2 SWS Lecture, 2 SWS Exercise

Additional Information

Dresden Database Research Group

Time

- Lecture: Thursday, 6.DS (16:40Uhr to 18:10Uhr) in SCH/A118
- Exercise: Wednesday, 4.DS (13:00Uhr to 14:30Uhr) in APB/E001/U
- Have a look on the OPAL page for
  - Lecture Notes
  - Updates and current Information
- Parallel stream on Zoom
  - Lecture: https://tu-dresden.zoom.us/j/69111351055?pwd=eGcvT3NwYWVKZGl5STljajFqVVRUQT09
  - Exercise: https://tu-dresden.zoom.us/j/68581716895?pwd=YXJXVFhEaWM3OGROQ1BiZWxVME9CUT09

Remarks and Questions

- Go for it!
- Questions are asked best during the lecture or e-mail us.

Exam

- Written exam for CMS.
- Oral exam for everyone else.

Time Schedule (tentative)

Dresden Database Research Group

October						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
					1	2
3	4	5	6	7	8	13
10	11	12	13	14	15	20
17	18	19	20	21	22	27
24	25	26	27	28	29	30
31						

November						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

December						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

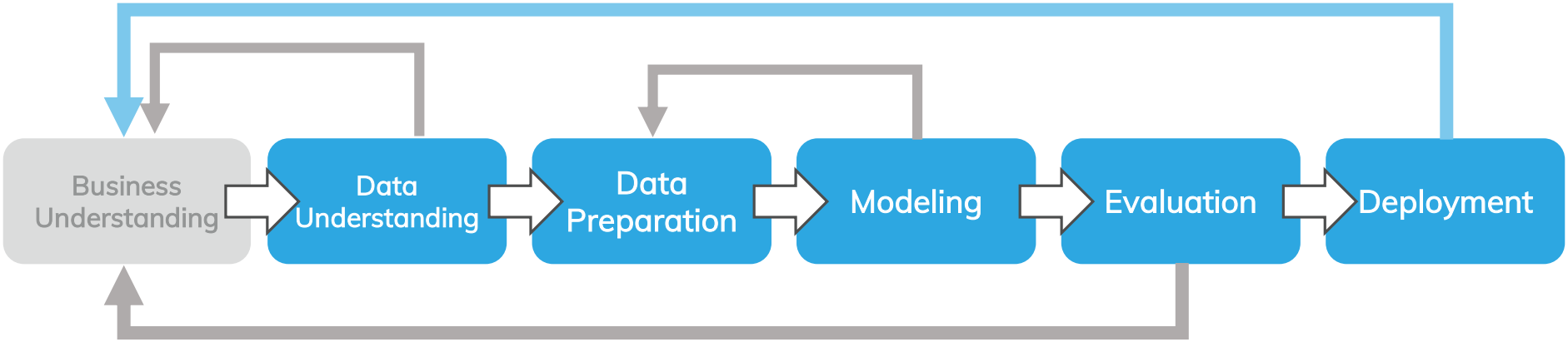
January						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1st Feb	2nd Feb			

- Today
- Lecture
- Exercise
- Free

Structure Data Analytics I

Dresden Database Research Group

Cross-Industry Standard Process for Data Mining (Crisp-DM)



Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
Business Objectives Assess Situation Data Mining Goals Project Plan	Collect Initial Data Describe Data Explore Data Verify Data Quality	Select Data Clean Data Integrate Data Format Data	Modeling Technique General Test Design Parameter Settings Assess Model	Evaluate Results Approve Models Next Steps	Plan Deployment Plan Monitoring Plan Maintenance Report and Review

Some Final Remarks

Dresden Database Research Group

Literature

- Will be recommended topic by topic

Copyright

- Commercial use of the lecture material outside of TUD is not allowed.